STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Cloning, expression, crystallization and crystallographic analysis of CouR from *Rhodopseudomonas palustris*

Chen Pan,[a,b] Yong-lin Hu,[b] Xiang-ning Jiang[a,c] and Ying Gai[a]*

[a]College of Biological Science and Technology, Beijing Forestry University, Beijing 100083, People's Republic of China, [b]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, People's Republic of China, and [c]The Tree and Ornamental Plant Breeding and Biotechnology Laboratory of Chinese Forestry Administration, Beijing 100083, People's Republic of China. *Correspondence e-mail: gaiying@bjfu.edu.cn

Received 23 August 2015
Accepted 8 October 2015

Edited by T. Bergfors, Uppsala University, Sweden

Keywords: MarR family; transcriptional regulator; protein–ligand complex; crystallographic study; lignin; *p*-coumarate.

CouR from *Rhodopseudomonas palustris* is a member of the MarR transcriptional regulator family. It regulates the expression of CouA and CouB, enzymes that are involved in the degradation of *p*-coumarate. *In vivo*, CouR binds to a DNA fragment containing the *couAB* promoter and suppresses the expression of CouA and CouB, while binding of *p*-coumaroyl-CoA attenuates its affinity towards DNA and activates the expression of CouA and CouB. Here, the crystallization and X-ray diffraction analyses of CouR alone and in complex with *p*-coumaroyl-CoA are reported. Apo and ligand-complexed CouR crystals diffracted to 2.5 and 3.3 Å resolution, respectively. The crystals of apo CouR belonged to space group $P22_12_1$, with unit-cell parameters $a = 62.78$, $b = 76.15$, $c = 87.38$ Å, whereas the crystals of the CouR–ligand complex belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 61.37$, $b = 69.82$, $c = 70.32$ Å. The crystals were predicted to contain two CouR molecules or CouR–ligand complexes per asymmetric unit.

1. Introduction

The abundance of aromatic compounds in nature, introduced either by human activities or from decaying plant materials, has led to the evolution of metabolic pathways in soil-dwelling bacteria to utilize these chemicals as a carbon source. Because of their importance in biodegradation, these metabolic pathways have been the focus of research and many pathways have been identified. One area of interest is how these pathways are regulated by the presence of different aromatic compounds. Several families of bacterial transcriptional regulators of the degradation pathways of aromatic compounds have been identified (Tropel & van der Meer, 2004).

p-Coumarate is produced in large quantities in green plants *via* the phenylpropanoid pathway. Most of this compound is converted into lignols, which are subsequently polymerized to form lignin, the second most abundant biopolymer on earth after cellulose. In addition to lignin, the biosynthesis of many compounds such as antimicrobials, molecular signals, aromatics, pigments and UV protectors requires *p*-coumarate as a precursor.

Several soil-dwelling bacteria have evolved metabolic pathways that can use lignin-derived compounds as a carbon source. In *Streptomyces coelicolor*, for example, the *pca* genes encode enzymes that are involved in the catabolism of protocatechuate and catechol, and these genes are negatively regulated by the transcriptional factor SCO6704 in the presence of lignin-derived aromatic compounds (Davis &

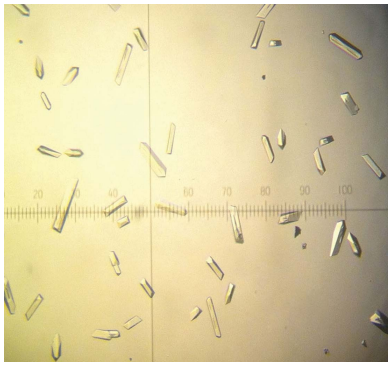


Table 1
Macromolecule-production information.

Source organism	*R. palustris*
DNA sequence	ATG ACC TCC TCC AAC CGT ATC ACC TCC CCG GCT ATG ACC GCT TCC AAA ACC GCT GCT GTT GCT AAA CCG ACC CGT GCT GGT CGT AAA GCT CCG GCT GTT GAA ACC GCT CCG GAA GCT TCC GAA CTG AAA ATG GGT GAA CTG TCC GAA CTG CTG GGT TAC GCT CTG AAA CGT GCT CAG CTG CGT GTT TTC GAA GAC TTC CTG CAC TGT GTT GCT CCG GTT CAG CTG ACC CCG GCT CAG TTC TCC GTT CTG CTG CTG CTG GAC GCT AAC CCG GGT CGT AAC CAG ACC GAA ATC GCT ACC ACC CTG GGT ATC CTG CGT CCG AAC TTC GTT GCT ATG CTG GAC GCT CTG GAA GGT CGT GGT CTG TGT GTT CGT ACC CGT TCC CCG TCC GAC CGT CGT TCC CAC ATC CTG ATG CTG ACC GAC AAA GGT CGT GCT ACC CTG GCT CGT GCT AAA AAA CTG GTT GCT ACC CGT CAC GAA GAC CGT CTG ACC GAA CTG CTG GGT CGT GAC AAC CGT GAC GCT CTG CTG TCC ATG CTG GCT ACC ATC GCT CGT GAA TTC TGA
Cloning vector	pET-28a
Expression vector	pET-28a
Expression host	*E. coli* strain BL21
Complete amino-acid sequence of the construct produced	MTSSNRITSPAMTASKTAAVAKPTRAGRKAPAVE-TAPEASELKMGELSELLGYALKRAQLRVFEDF-LHCVAPVQLTPAQFSVLLLLDANPGRNQTEIA-TTLGILRPNFVAMLDALEGRGLCVRTRSPSDR-RSHILMLTDKGRATLARAKKLVATRHEDRLTE-LLGRDNRDALLSMLATIAREF

Sello, 2010). Similar systems are also found in *Agrobacterium tumefaciens* (Parke, 1995), *Corynebacterium glutamicum* (Brinkrolf *et al.*, 2006) and *Acinetobacter baylyi* (Siehler *et al.*, 2007). Nevertheless, research on proteins targeting *p*-coumarate, the starting compound of lignin biosynthesis, has been scarce.

In a recent study, it has been found that *Rhodopseudomonas palustris*, a Gram-negative purple nonsulfur bacterium, can utilize *p*-coumarate as a carbon source. The enoyl-CoA lyase CouB converts *p*-coumarate into *p*-coumaroyl-CoA, and the enoyl-CoA hydratase CouA cleaves the ester side chain of *p*-coumaroyl-CoA to form acetyl-CoA and benzaldehyde (Hirakawa *et al.*, 2012). CouR, a multiple antibiotic-resistance regulator (MarR) repressor family protein, regulates the gene expression of CouA and CouB in response to *p*-coumaroyl-CoA. The *couR* gene is located downstream of the *couAB* genes, and the CouR protein binds a 29 bp region upstream of *couA* and protects it from DNase I digestion. This region encompasses an inverted repeat sequence separated by a 3 bp spacer and overlaps with the −10 promoter element. The study also found that the binding between CouR and the protected region was not abolished by either *p*-coumarate or other degradation products, but only by *p*-coumaroyl-CoA (Hirakawa *et al.*, 2012). A homologous transcriptional regulator has also been identified in *Sphingobium* sp. strain SYK-6 (Kasai *et al.*, 2012), in which the MarR-family protein FerC negatively regulates the expression of FerBA, with feruloyl-CoA, *p*-coumaroyl-CoA, caffeoyl-CoA and sinapoyl-CoA as inducers. However, the molecular mechanism explaining how these transcriptional regulators interact with *p*-coumaroyl-CoA or its derivatives and modulate their DNA-binding affinities accordingly has remained unresolved (Perera & Grove, 2010).

Our research groups are interested in the biosynthesis and degradation of lignin and the phenylpropanoid compounds of the pathway, so we embarked on structural analysis of CouR in order to illustrate how this protein recognizes and binds *p*-coumaroyl-CoA and the structural basis for the attenuated DNA binding of CouR upon its interaction with *p*-coumaroyl-CoA.

Here, we report the expression, purification and crystallization of CouR and its complex with *p*-coumarate. We also present the crystallographic results of our X-ray diffraction experiments on these two crystal forms.

2. Materials and methods

2.1. Cloning, expression and purification of CouR

The gene encoding CouR was redesigned to optimize protein expression in *Escherichia coli*. It was synthesized by Sangon Biotech Company, Shanghai, People's Republic of China and was ligated to a pET-28a vector, which provides an N-terminal His_6 tag on the expressed protein. The plasmid was transformed into *E. coli* strain BL21 using the heat-shock method (Table 1). The transformed bacteria were grown at 310 K in the presence of 25 mg l^{-1} kanamycin until the OD_{600} reached 0.6–0.8. Protein overexpression was induced by adding 0.05 m*M* isopropyl β-D-1-thiogalactopyranoside and incubating at 295 K. After 12 h of incubation, the cells were harvested by centrifugation at 4000*g* for 30 min at 277 K.

The bacteria were lysed by sonication on ice at 200 W using 4 s pulses with 9 s intervals for 99 cycles. Insoluble debris was removed by centrifugation at 16 000*g* for 40 min at 277 K. The supernatant was loaded onto Ni^{2+}–NTA His·Bind Superflow resin (Novagen, Darmstadt, Germany) pre-equilibrated with lysis buffer (50 m*M* Tris–HCl pH 7.5, 300 m*M* NaCl, 10 m*M* imidazole). After washing the resin with wash buffer 1 (50 m*M* Tris–HCl pH 7.5, 1 *M* NaCl, 20 m*M* imidazole) and wash buffer 2 (50 m*M* Tris–HCl pH 7.5, 1 *M* NaCl, 50 m*M* imidazole), the His_6-tagged protein was eluted with an elution buffer consisting of 42 m*M* Na_2HPO_4, 8 m*M* NaH_2PO_4, 300 m*M* NaCl, 250 m*M* imidazole followed by buffer exchange on a 5 ml HiTrap Desalting column (GE Healthcare, Uppsala, Sweden) with storage buffer (20 m*M* Tris–HCl pH 7.5, 50 m*M* KCl, 1 m*M* DTT, 8% glycerol). Purified CouR was concentrated by centrifugation using an Amicon 30 kDa centrifugal concentrator (Millipore) and the protein concentration was measured by the absorbance at 280 nm. The protein was stored in the storage buffer and frozen at 193 K until further use.

p-Coumaroyl-CoA was enzymatically synthesized and purified using an improved method based on the work of Till Beuerle (Beuerle & Pichersky, 2002), after which it was lyophilized. The compound was dissolved in distilled

Table 2
Crystallization.

	CouR	CouR–ligand
Final method	Hanging-drop vapour diffusion	Hanging-drop vapour diffusion
Plate type	16-well protein crystallization plates	16-well protein crystallization plates
Temperature (K)	293	293
Protein concentration (mg ml^{-1})	35	80
Buffer composition of protein solution	20 m*M* Tris–HCl pH 7.5, 50 m*M* KCl, 1 m*M* DTT, 8% glycerol	20 m*M* Tris–HCl pH 7.5, 50 m*M* KCl, 1 m*M* DTT, 8% glycerol
Composition of reservoir solution	0.1 *M* Tris–HCl pH 8.0, 3.0 *M* sodium formate	0.1 *M* bis-tris–HCl pH 6.5, 2%(*w*/*v*) Tacsimate pH 6.0, 30%(*w*/*v*) PEG 3350
Volume and ratio of drop	1 µl drops, 1:1 ratio	1 µl drops, 1:1 ratio
Volume of reservoir (µl)	500	500

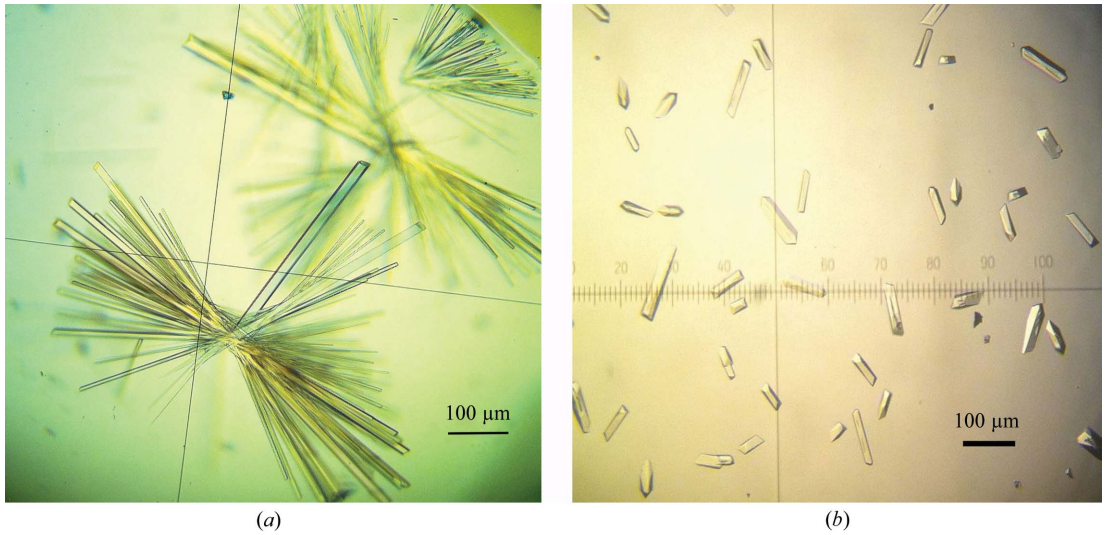
deionized water and its concentration was estimated spectrophotometrically before the co-crystallization experiments.

2.2. Crystallization

The preliminary crystallization conditions for CouR and the CouR–ligand complex were screened by the sitting-drop vapour-diffusion method at 293 K using commercial crystallization kits from Hampton Research, Aliso Viejo, California, USA. Each crystallization drop consisted of 1 µl protein solution (20 mg ml^{-1} CouR) mixed with 1 µl screen solution and was equilibrated against 0.1 ml reservoir solution. During optimization of the crystallization conditions, the method was changed to hanging-drop vapour diffusion. For apo CouR, preliminary crystals were observed in a condition consisting of 0.1 *M* Tris–HCl pH 8.0, 3.5 *M* sodium formate. By reducing the concentration of sodium formate to 3.0 *M* and increasing the concentration of CouR to 35 mg ml^{-1}, we obtained crystals that were suitable for diffraction data collection (Fig. 1*a*). For crystallization of the CouR–ligand complex, a molar ratio of 1:5 of CouR:ligand was used. Initial complex crystals were obtained in a condition consisting of 0.1 *M* bis-tris pH 6.5, 2%(*v*/*v*) Tacsimate pH 6.0, 20%(*w*/*v*) PEG 3350. To obtain crystals that were suitable for data collection, we increased the concentration of PEG 3350 to 30% and the concentration of CouR to 80 mg ml^{-1} while maintaining the molar ratio of CouR and its ligand. We also added DNase I (Bio Basic Inc., Markham, Canada) at a molar ratio of 1:1000 (DNase I:CouR) to the protein–ligand solution. The inclusion of DNase I was necessary to optimize the crystals of the protein–ligand complex (Fig. 1*b*). When preparing the protein–ligand solution, we first mixed CouR with DNase I, followed by addition of *p*-coumarate-CoA. Precipitation resulted when the ligand was added and subsequently cleared on incubation on ice for 3 h. Detailed crystallization information is summarized in Table 2.

2.3. Data collection and processing

The CouR crystals were directly mounted on nylon CryoLoops (Hampton Research) and flash-cooled in a nitrogen stream at 100 K. In contrast, the crystals of the CouR–ligand complex were first soaked for several seconds in a cryoprotectant solution that consisted of 95% reservoir solution and 5% glycerol. They were then transferred into paraffin oil using nylon CryoLoops for approximately 5 s before being flash-cooled in a nitrogen stream. Diffraction data were collected at a wavelength of 0.9793 Å on beamline BL17U-MX at



(*a*) (*b*)

Figure 1
Crystals of CouR and the CouR–ligand complex. (*a*) Crystals of CouR. The largest is approximately 400 ×14 ×14 µm in size. (*b*) Crystals of the CouR–ligand complex.

Shanghai Synchrotron Radiation Facility (SSRF) with a MAR DTB detector (Fig. 2). All diffraction data sets were processed using *MOSFLM* (Leslie, 2006) and were scaled using *SCALA* (Evans, 2006) from the *CCP*4 suite (Winn *et al.*, 2011). Data-collection statistics are listed in Table 3.

3. Results and discussion

The DNA encoding CouR was optimized for *E. coli* expression and a yield of approximately 27 mg protein per litre of overexpression medium was observed in our experiments. During the Ni–NTA affinity purification, a concentration of 1.0 *M* NaCl was required in the presence of 20 and 50 m*M* imidazole to keep the target protein bound while removing other contaminating molecules from the column. After the Ni–NTA affinity column, we used a desalting column to remove the NaCl and imidazole and to change the buffer condition to one that was more suitable for crystallization screening. The purity of the CouR sample after the desalting column was greater than 95% as assayed by SDS–PAGE (Fig. 3).

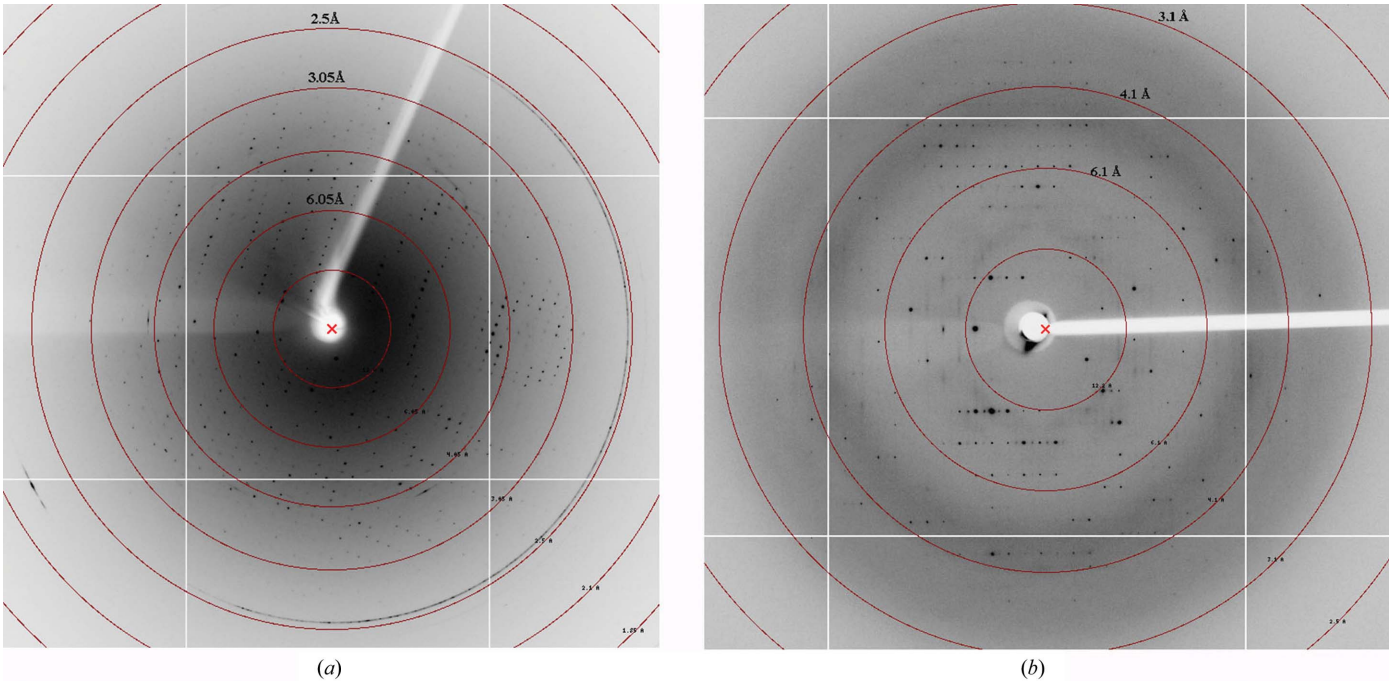
The initial crystals of apo CouR obtained at a protein concentration of 20 mg ml^{-1} were clustered thin needles and were difficult to optimize, despite efforts with a wide-ranging screen of additives and varying other conditions such as the pH and the precipitant concentrations. After we increased the concentration of CouR to 35 mg ml^{-1}, however, we were able to obtain crystals with a thicker needle shape. The crystals used in data collection were cut off from the needle clusters.

The crystallization behaviour of CouR changed dramatically in the presence of its ligand. Protein crystals were only observed in conditions containing PEG 3350, from which apo CouR had failed to crystallize (Table 2). Based on these observations and the fact that the crystals obtained in the presence of the CouR ligand belonged to a different space group with different unit-cell parameters from those of the apo CouR crystals, we believed that the crystals were of the CouR–ligand complex rather than apo CouR crystals with a different morphology. For crystallization of the CouR–ligand complex, increasing the protein concentration to 80 mg ml^{-1}

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

	CouR	CouR–ligand
Diffraction source	Beamline BL17U, SSRF	Beamline BL17U, SSRF
Wavelength (Å)	0.97930	0.97930
Temperature (K)	100	100
Detector	MAR DTB	MAR DTB
Crystal-to-detector distance (mm)	280	280
Rotation range per image (°)	1	1
Total rotation range (°)	180	180
Exposure time per image (s)	1	1
Space group	$P22_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 62.78, b = 76.15, c = 87.38	a = 61.37, b = 69.82, c = 70.32
Mosaicity (°)	0.68	0.85
Resolution range (Å)	48.44–2.50 (2.64–2.50)	49.55–3.30 (3.48–3.30)
Total No. of reflections	88646 (12416)	17842 (2615)
No. of unique reflections	15044 (2174)	4626 (662)
Completeness (%)	99.8 (100)	96.0 (97.0)
Multiplicity	5.9 (5.7)	3.9 (4.0)
$\langle I/\sigma(I)\rangle$	9.4 (2.8)	8.0 (2.0)
$R_{\rm r.i.m.}$	0.052 (0.327)	0.059 (0.499)
Overall B factor from Wilson plot (Å^2)	40.9	87.9



(*a*) (*b*)

Figure 2
Representative diffraction images. (*a*) Diffraction image of an apo CouR crystal. (*b*) Diffraction image of a CouR–ligand complex crystal.

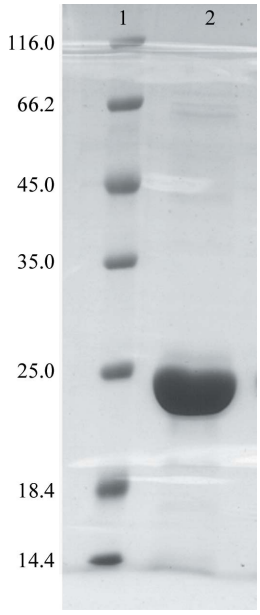


Figure 3
12% SDS–PAGE. Lane 1, protein markers (molecular masses are labelled in kDa). Lane 2, purified CouR.

also helped to obtain larger crystals, but the crystals were polycrystalline and were of poor quality; integration of the diffraction frames was unsuccessful. After many attempts, we eventually found that the inclusion of DNase I in the protein–ligand mixture at a molar ratio of 1:1000 (DNase I:CouR) led to single crystals with an improved diffraction quality. We propose that DNase I degrades the trace amount of DNA that binds to CouR and improves the homogeneity of the protein.

To determine the structures of CouR and the CouR–ligand complex, we plan to solve the phase problem by the molecular-replacement method using homologous MarR protein structures, such as PDB entries 2pex (Newberry *et al.*, 2007) and 3bpv (Saridakis *et al.*, 2008), as initial models. The results reported here lay a solid foundation for further research to reveal the molecular mechanisms of how CouR and other homologous transcriptional regulators recognize and bind *p*-coumaroyl-CoA and modulate its DNA-binding affinities.

Acknowledgements

We are grateful for the technical support of the staff at beamline BL17U of Shanghai Synchrotron Radiation Facility during data collection. This work was supported by grants from the Project of the National Natural Science Foundation of China (No. 31300498) to YG and from the National High Technology Research and Development (863 Program, 2011AA100203) to XNJ. This work was carried out in the Tree and Ornamental Plant Breeding and Biotechnology Laboratory of the State Forestry Administration and the National Laboratory of Protein Engineering.

References

Beuerle, T. & Pichersky, E. (2002). *Anal. Biochem.* **302**, 305–312.
Brinkrolf, K., Brune, I. & Tauch, A. (2006). *Genet. Mol. Res.* **5**, 773–789.
Davis, J. R. & Sello, J. K. (2010). *Appl. Microbiol. Biotechnol.* **86**, 921–929.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Hirakawa, H., Schaefer, A. L., Greenberg, E. P. & Harwood, C. S. (2012). *J. Bacteriol.* **194**, 1960–1967.
Kasai, D., Kamimura, N., Tani, K., Umeda, S., Abe, T., Fukuda, M. & Masai, E. (2012). *FEMS Microbiol. Lett.* **332**, 68–75.
Leslie, A. G. W. (2006). *Acta Cryst.* D**62**, 48–57.
Newberry, K. J., Fuangthong, M., Panmanee, W., Mongkolsuk, S. & Brennan, R. G. (2007). *Mol. Cell*, **28**, 652–664.
Parke, D. (1995). *J. Bacteriol.* **177**, 3808–3817.
Perera, I. C. & Grove, A. (2010). *J. Mol. Cell. Biol.* **2**, 243–254.
Saridakis, V., Shahinas, D., Xu, X. & Christendat, D. (2008). *J. Mol. Biol.* **377**, 655–667.
Siehler, S. Y., Dal, S., Fischer, R., Patz, P. & Gerischer, U. (2007). *Appl. Environ. Microbiol.* **73**, 232–242.
Tropel, D. & van der Meer, J. R. (2004). *Microbiol. Mol. Biol. Rev.* **68**, 474–500.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.